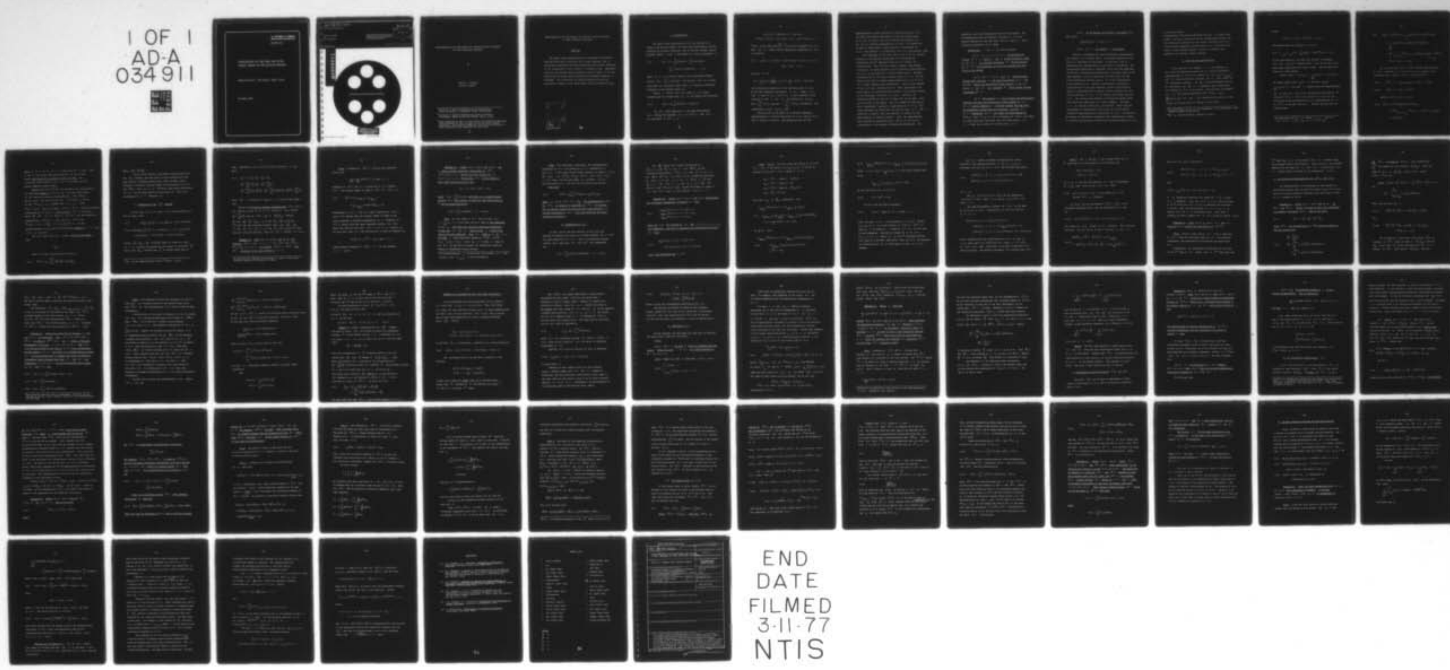
AD-A034 911 BROWN UNIV PROVIDENCE R I LEFSCHETZ CENTER FOR DYNAM--ETC F/G 12/1
APPROXIMATIONS FOR FUNCTIONALS AND OPTIMAL CONTROL PROBLEMS ON --ETC(U)
NOV 76 H J KUSHNER, G DIMASI N00014-76-C-0279
UNCLASSIFIED LCDS-TR-76-2 AFOSR-TR-77-0018 NL

1 OF 1
AD-A
034911



END
DATE
FILMED
3-11-77
NTIS

U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

AD-A034 911

APPROXIMATIONS FOR FUNCTIONALS AND OPTIMAL CONTROL PROBLEMS ON JUMP DIFFUSION PROCESSES

BROWN UNIVERSITY, PROVIDENCE, RHODE ISLAND

NOVEMBER 1976

AFOSR - TR - 77 - 0018

LCDS TECHNICAL REPORT 76-2

HAROLD J. KUSHNER
and
GIOVANNI DIMASI

APPROXIMATIONS FOR FUNCTIONALS AND OPTIMAL CONTROL PROBLEMS ON JUMP DIFFUSION PROCESSES

NOVEMBER 1976

LEFSCHETZ CENTER FOR DYNAMICAL SYSTEMS

DIVISION OF APPLIED MATHEMATICS

BROWN UNIVERSITY

APPROXIMATIONS FOR FUNCTIONALS AND OPTIMAL CONTROL PROBLEMS ON JUMP DIFFUSION PROCESSES*

by

Harold J. Kushner+
Giovanni DiMasi++

+Division of Applied Mathematics and Engineering, Brown University, Providence, Rhode Island 02912

++Division of Applied Mathematics, Brown University, Providence, Rhode Island and LADSEB, Padua, Italy.

*Work supported by the Air Force Office of Scientific Research, under AF-AFOSR 76-3063, in part by the National Science Foundation, under Eng-73-03846-A01, and in part by Office of Naval Research, under NONR N000 14-76-C-0279.

Approximations for Functionals and Optimal Control Problems on Jump Diffusion Processes

Abstract

The paper treats approximations to stochastic differential equations with both a diffusion and a jump component, and to associated functionals and partial-differential-integral equations of the (degenerate or not) elliptic or parabolic type. Approximations for the optimal control problem on such a model, or for the associated non-linear partial-differential-integral equation is discussed. The techniques are purely probabilistic and are extensions of those in [3], which dealt with the diffusion case.

1. Introduction

The paper treats approximations for the process (1.1) (terms to be defined below), and associated functionals, partial differential integral equations, and for certain optimal control problems (where f and k can depend on a control parameter u).

$$(1.1)\qquad X(t) = x + \int_0^t f(X(s))ds + \int_0^t \sigma(X(s))dw(s)$$

$$+ \int_0^t \int g(X(s^-),\alpha)Q(d\alpha\times ds), \quad t \geq 0,$$

where $Q(\cdot,\cdot)$ is a Poisson measure, with associated Poisson process $Q(t) = \int_0^t \int \alpha Q(d\alpha\times ds)$. The process (1.1) is a widely used model for situations where there is a jump and a diffusion component to the process paths [1].

Define $\tau = \inf\{t: X(t) \notin G\}$, where G is a given bounded open set, assume that $E_x\tau < \infty$, and define the functional

$$(1.2)\qquad R(x) = E_x \int_0^\tau k(X(s))ds + E_x\phi(X(\tau)).$$

Let $Q(\cdot)$ have jump rate c, and jump distribution $p(\cdot)$. Define the measures $\pi(\cdot,\cdot),\pi(\cdot),\hat{\Gamma}(x,\cdot)$ and $\Gamma(x,\cdot)$ by (for Borel A and $t > s$)

$$\pi(A,[s,t]) = EQ(A\times[s,t]) = \pi(A)(t-s),$$

$$\hat{\Gamma}(x,A) = c\Gamma(x,A) = \pi(\alpha: g(x,\alpha) \in A) = cp(\alpha: g(x,\alpha) \in A).$$

$\hat{\Gamma}(x,A)$ is the jump rate into A at t of the last integral in (1.1), when $X(t) = x$. Under certain smoothness assumptions on $R(\cdot)$, it satisfies

$$\text{(1.3)} \qquad \mathcal{L}V(x) + \int [V(x+\alpha) - V(x)]\hat{\Gamma}(x,d\alpha) + k(x) = 0,\ x \in G,$$

$$V(x) = \phi(x),\quad x \notin G,$$

the term $\mathcal{L}$ is

$$\mathcal{L} = \sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j f_i(x)\frac{\partial}{\partial x_i},\ 2a(x) = \sigma(x)\sigma'(x),$$

the differential generator of the diffusion part of (1.1). We use the following conventions. If $Q(\cdot)$ jumps α at time t, then in (1.1), $g(X(t^-),\alpha)$ is the increment of the integral at time t. Let $Y(\cdot)$ be constant on $[t_1,t_2)$. Then $\int_{t_1}^{t_2}\int g(Y(s^-),\alpha)Q(d\alpha\times ds) = \int_{t_1}^{t_2}\int g(Y(t_1),\alpha)Q(d\alpha\times ds)$. The integration is over $(t_1,t_2]$ only.

The main aim of the paper is to develop computable approximations to various functionals of (1.1), such as (1.2), with or without a control. The approximations are also

approximations to weak solutions of equations such as (1.3) (or their non-linear counterparts in the controlled case). As in [2], we exploit the close relationship between (1.2) and (1.3) to develop approximations to both. The method can be used to approximate a broad variety of functionals of (1.1). (See Theorem 7.1.) Functional (1.2) is a very special case.

The technique is like that used in [2] or in [3] (we use results in [3] whenever convenient) for functionals of diffusions, and for equations such as (1.3) without the integral term. See, also [4], where a simpler jump problem is treated. By using suitable finite difference approximations to (1.3), we can construct a certain Markov chain. Suitable interpolations of the chain converge weakly to (1.1), and the chain can be used to compute approximations to (1.2) or (1.3) (or to other suitable path functionals). This is true whether or not (1.3) has a smooth solution. Under conditions to be given below, the approximations converge to the correct values as the difference intervals go to zero. Actually, the difference method is only one of many that can be used - this will be clear in the sequel. The f.d. method is used only to get a "consistent" sequence of approximations to $X(\cdot)$. Any other technique for doing this can also be used. If f and k depend on a control, then the technique is useful for the approximation of optimal controls. Then, the approximating chain becomes a controlled chain. See [2] or [3] for an introduction to the general technique and background. For

simplicity, only the homogeneous case will be treated. The non-homogeneous case and parabolic versions of (1.3) are treated similarly to the case here (see [3], Chap. 7). Also, reflecting boundaries can be added.

Assumptions. r and r' are given integers

(A1.1) $f(\cdot)$, $\sigma(\cdot)$, $k(\cdot)$, $\phi(\cdot)$ are bounded continuous R^r, $r \times r$ matrix, R and R valued functions, resp., on R^r; $g(\cdot,\cdot)$ is a bounded measurable R^r valued function on $R^r \times R^{r'}$, and continuous in its first argument for each value of the second.

(A1.2) $Q(t)$, $t \in [0,\infty)$ is a $R^{r'}$ valued Poisson process with jump rate c, and jump distribution $p(\cdot)$. Let $Q(d\alpha \times ds)$ denote the associated Poisson measure [1] or [5], Chapter 6. Let $w(\cdot)$ be a standard R^r valued Wiener process independent of $Q(\cdot)$.

(A1.3) The process (1.1) has a unique non-anticipative solution, for each non-anticipative (with respect to $w(\cdot)$, $Q(\cdot,\cdot)$) initial condition x, with and without the jump term. By uniqueness, we mean that the solutions - for any $w(\cdot)$, $Q(\cdot,\cdot)$ satisfying (A1.2) - all induce the same measure on $D^r[0,\infty)$ (see [3] or [6] for a discussion of $D[0,\infty)$, the space of right continuous functions with left hand limits). If $c < \infty$, then (A1.3) holds if it holds with $g \equiv 0$.

(A1.4) <u>(to be dropped in Section 8; see below</u> (2.1) <u>for</u> $Q_h(x)$)

$$\sup_x (h^2/Q_h(x)) \to 0 \quad \underline{\text{as}} \quad h \to 0.$$

(A1.5) $E_x \tau < \infty$ <u>for values</u> x <u>of interest</u>.

Section 2 introduces the finite difference approximation and relates it to a Markov chain. The finite difference solution is a functional of the chain. In Sections 2 and 3, a continuous time interpolation of the chain is introduced. The interpolated process has a diffusion, drift and jump component, and Section 4 discusses the properties of the weak limit of the jump component, and also shows that the weak limit (as the finite difference interval goes to 0) of the interpolation is the process (1.1). Section 5 discusses an alternative representation of the jump component, which is particularly useful in approximations to optimal control problems for controlled versions of (1.1), or for certain non-linear forms of (1.3). The section also contains an alternative Markov chain with which the computation is a little simpler. The case $c = \infty$ is treated in Section 6, and Section 7 deals with the convergence of functionals of the chain (or of the finite difference approximation) to functionals of (1.1). In Section 8, we develop an alternative continuous time interpolation, which is a Markov jump process and which also converges weakly to (1.1).

As with the earlier interpolation, the finite difference solution - or chain functional - is also a functional of the interpolated process, and this functional converges to the correct functional of (1.1), as the finite difference interval goes to 0. The limit functional is the desired solution. Section 9 contains some remarks on the optimal control problem.

2. The Discrete Approximation

Until Section 6, we assume $c < \infty$. The operator in (1.3) will be discretized using the f.d. (finite difference) approximations of [2] or [3], Chapter 6.2, for $V_{x_i}(\cdot), V_{x_i x_j}(\cdot)$. Let[+] h = finite difference interval, and e_i = unit vector in i^{th} coordinate direction. Let R_h^r = finite difference grid on R^r, and define $G_h = R_h^r \cap G$. A convenient way to discretize the integral in (1.3) is as follows. For each set of integers[++] $j_1, \ldots, j_r$ and "finite difference box", $b^h(j_1, \ldots, j_r) = \prod_{i=1}^{r} (j_i h - h, j_i h]$, let $j_1' h, \ldots, j_r' h$ denote the grid point in the closure of the box which is closest to the origin.

[+]For simplicity, we let h be independent of the direction. More general schemes are possible.

[++]The j_i can be positive, negative or zero.

Define

$$\Gamma^h(x,j_1,\ldots,j_r) = \Gamma(x,b^h(j_1,\ldots,j_r)).$$

Then approximate the integral in (1.3) by

$$(2.1) \quad c \sum_{j_1,\ldots,j_r} [V(x + \sum_{i=1}^{r} e_i j_i' h) - V(x)]\Gamma^h(x,j_1,\ldots,j_r).$$

By the approximation, any jump into the box is remapped into $j_1',\ldots,j_r'$, the "closest" point in the box to the origin. Many other conventions will work as well. The one above was chosen for definiteness.

Define $Q_h(x) = 2 \sum_i a_{ii}(x) - \sum_{\substack{i,j \\ i\neq j}} |a_{ij}(x)| + h \sum_i |f_i(x)|$.

We always suppose that $Q_h(x) > 0$, and that $a_{ii}(x) - \sum_{\substack{i\neq j \\ j}} |a_{ij}(x)| \geq 0$ for each i. Substituting the approximations for $\mathscr{L}$, and (2.1), into (1.3), multiplying each term by h^2, defining[+] $p^h(\cdot,\cdot)$ as in [3], Chapter 6.2, or as in [2], collecting terms and denoting the f.d. approximation to $V(\cdot)$ by $V^h(\cdot)$, we get (the sum over $\pm$ denotes the sum over all combinations)

[+]In particular, $p^h(x,y) = 0$ unless $y = x \pm e_i h$, or (for $i \neq j$) $x \pm e_i h \pm e_j h$ or $x \pm e_i h \mp e_j h$.

$$(2.2) \qquad [Q_h(x) + ch^2]V^h(x) = \sum_{\substack{i,j\\ \pm}} Q_h(x)p^h(x,x\pm e_i h)V^h(x\pm e_i h)$$

$$+ \sum_{\substack{\pm\\ i,j\\ i\neq j}} Q_h(x)p^h(x,x\pm e_i h\mp e_j h)V^h(x\pm e_i h\mp e_j h)$$

$$+ ch^2 \sum_{j_1,\ldots,j_r} \Gamma^h(x,j_1,\ldots,j_r)V^h(x+\sum_i e_i j_i' h) + h^2 k(x),$$

$$x \in G_h,$$

$$V^h(x) = \phi(x), \quad x \notin G_h.$$

We can work with (2.2), or with various approximations to (2.2). Equation (2.2) will be put into a slightly more convenient form. Define $\Delta t^h(x) = h^2/[Q_h(x)+ch^2]$.

$$(2.3a) \qquad \hat{P}^h(x) = 1 - (Q_h(x)/[Q_h(x) + ch^2]) = c\Delta t^h(x)$$

$$(2.3b) \qquad \tilde{P}^h(x) = 1 - \exp - c\Delta t^h(x).$$

Rewrite (2.2) as (for $x \in G_h$)

$$(2.4) \qquad V^h(x) = (1-P^h(x))[\sum_y p^h(x,y)V^h(x,y)]$$

$$+ P^h(x)[\sum_{j_1,\ldots,j_r} \Gamma^h(x,j_1,\ldots,j_r)V^h(x + \sum_i e_i j_i' h]$$

$$+ \Delta t^h(x)k(x),$$

where P is $\hat{P}$ or $\tilde{P}$. If $\hat{P}$ is used, then (2.4) = (2.2). The $\hat{P}$ and $\tilde{P}$ are first order approximations (in Δt) to each other; the limits do not depend on which is used. For notational definiteness we use (2.3b) (which seems to give better numerical results also).

Suppose that the Γ^h, P^h, p^h are defined on all grid points x (or grid point pairs for p^h) on R_h^r. Then the coefficients in (2.4) have the following interpretation. The $p^h(x,y)$ are in $[0,1]$ and sum to unity over y, for each x, and so do the $\Gamma^h(x,j_1,\dots,j_r)$ (summed over $j_1,\dots,j_r$). Also, $P^h(x) \varepsilon [0,1]$. Thus, the system $\{p^h(x,y), P^h(x)\}$ has the following interpretation in terms of a Markov chain $\{\xi_n^h\}$ on the state space R_h^r. Let $\xi_n^h = x$. Then w.p. $(1-P^h(x))$, use the transition probabilities $\{p^h(x,y), y \varepsilon R_h^r\}$, and w.p. $P^h(x)$, use $\Gamma^h(x,j_1,\dots,j_r) \equiv P\{\xi_{n+1}^h = x + \sum_i e_i j_i' h | \xi_n^h = x\}$. In the second instance, we say that the process has jumped at $n+1$ - even if all j_i' are zero.

Let $N_h = \min\{n: \xi_n^h \not\in G_h\}$. Assume, only for the moment, that

$$E_x N_h < \infty.$$

Then (2.4) has a unique solution which is

$$(2.5) \qquad V^h(x) = E_x \sum_{n=0}^{N_h-1} k(\xi_n^h)\Delta t_n^h + E_x \phi(\xi_{N_h}^h),$$

where $\Delta t_n^h = \Delta t^h(\xi_n^h)$.

We will next develop a convenient representation for the $\{\xi_n^h\}$ process, and then prove various properties of the weak limits, ultimately showing that (when suitably interpolated), the chains converge weakly to the solution to (1.1), and $V^h(x) \to R(x)$, as $h \to 0$. Similarly, many other functionals of $X(\cdot)$ can be approximated by functionals of $\{\xi_n^h\}$, with convergence as $h \to 0$ (Theorem 7.1).

3. Properties of the $\{\xi_n^h\}$ Process

We have (see [2] or [3], Chap. 6, for the calculation in the "no jump" case),

$$(3.1) \qquad E[\xi_{n+1}^h - \xi_n^h | \ \xi_n^h = x, \text{ no jump at } n+1] = f(x)\Delta t^h(x)$$

$$(3.2)^+ \quad \text{Covar}[\xi_{n+1}^h - \xi_n^h | \ \xi_n^h = x, \text{ no jump at } n+1] = \textstyle\sum_h(x)\Delta t^h(x)$$

$$= 2a(x)\Delta t^h(x) + \Delta t^h(x)[h\tilde{f}(x) - f(x)f'(x)\Delta t^h(x)]$$

Define $\beta_n^h = \xi_{n+1}^h - \xi_n^h - f(\xi_n^h)\Delta t_n^h$, where if there is a jump at $n+1$, we alter the definition and assume (to calculate β_n^h only) that ξ_{n+1}^h evolves from ξ_n^h as though there were no

---

$^+\tilde{f}(x)$ is the diagonal matrix with i^{th} entry $|f_i(x)|$.

jump. Henceforth, let us fix the initial condition $x = \xi_0^h$. Write

$$(3.3) \qquad \xi_n^h = x + F_n^h + B_n^h + J_n^h - E_n^h,$$

$$F_n^h = \sum_{i=0}^{n-1} f(\xi_i^h)\Delta t_i^h, \quad B_n^h = \sum_{i=0}^{n-1} \beta_i^h,$$

$$J_n^h = \sum_{i=0}^{n-1} [\xi_{i+1}^h - \xi_i^h] I_i^h, \quad E_n^h = \sum_{i=0}^{n-1} [f(\xi_i^h)\Delta t_i^h + \beta_i^h] I_i^h \equiv \sum_{i=0}^{n-1} \varepsilon_i^h,$$

where $I_i^h = 1$ if there is a jump at $i + 1$, and is zero otherwise.

Define the <u>piecewise constant interpolations</u> $\xi^h(\cdot)$, $F^h(\cdot)$, etc. exactly as done in [2] or [3], Chapter 6. E.g., define $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$, and set $F^h(t) = F_n^h$ on $[t_n^h, t_{n+1}^h)$. Define m_i^h and $\delta m_i^h = m_i^h - m_{i-1}^h$, $i \geq 2$, and $m_1^h = \delta m_1^h$, to be the i^{th} jump time and interjump interval resp., for $\{\xi_n^h\}$. Let τ_i^h and $\delta\tau_i^h = \tau_i^h - \tau_{i-1}^h$, $i \geq 2$, and $\tau_1^h = \delta\tau_1^h$ be the i^{th} jump time and interjump interval, resp. for $\xi^h(\cdot)$. Define $R^+ = [0,\infty)$.

<u>Theorem 3.1</u>. <u>Under</u> (A1.1), (A1.2), <u>and</u> (A1.4), <u>the sequence</u> $\{\Phi^h(\cdot), \delta\tau_i^h, i = 1,\ldots\} \equiv \{\xi^h(\cdot), F^h(\cdot), B^h(\cdot), J^h(\cdot), E^h(\cdot), \delta\tau_i^h, i = 1,2,\ldots\}$ <u>is tight[+] on</u> $D^{5r}[0,\infty) \times (R^+)^\infty$. <u>Also</u> $E^h(\cdot)$ <u>converges to the zero process as</u> $h \to 0$.

[+]We say that the sequence of processes is tight if the corresponding sequence of measures is tight.

<u>Proof</u>. Tightness of $\{E^h(\cdot)\}$ and the last assertion follow from

$$\lim_{h\to 0} E \max_{t\le T} |E^h(t)|^2 = 0, \text{ each } T < \infty.$$

Tightness of $\{F^h(\cdot), B^h(\cdot)\}$ is proved as in [3], Theorem 6.3.1. The average number of jumps of $J^h(\cdot)$ on $[0,t]$ is

$$(3.4) \qquad A_1^h = E \sum_i I_{\{\text{jump at } i\}} I_{\{t_i^h \le t\}} =$$

$$E \sum_i I_{\{t_i^h \le t\}} (1 - \exp - c\Delta t_{i-1}^h) \le ct.$$

Boundedness of $g(\cdot,\cdot)$ and (3.4) imply tightness of $\{J^h(\cdot)\}$, since (3.4) implies that the average number of jumps on any finite interval is bounded independently of h. Tightness of $\{\xi^h(\cdot)\}$ follows from the above tightness results. Let $\theta(h)$ denote any function such that $\theta(h) \to 0$ as $h \to 0$. Finally, we can show that (see, e.g., the calculations in proof of Theorem 4.1)

$$P\{\delta\tau_i^h \ge T\} \le e^{-cT} + \theta(h), \text{ each } T < \infty,$$

which implies tightness of $\{\delta\tau_i^h, i \ge 1\}$ for any sequence $h \to 0$. Q.E.D.

Theorem 3.2. Assume (A1.1), (A1.2) and (A1.4). Let h index a weakly convergent subsequence of $\{\hat{\Phi}^h\} \equiv \{\phi^h(\cdot), \delta\tau_i^h, i \geq 1\}$. Denote the limit process by $\hat{\Phi} = \{\phi(\cdot), \delta\tau_i, i \geq 1\}$. The process $\phi(\cdot)$ can be chosen to have right continuous paths, also

$$\xi(t) = x + F(t) + B(t) + J(t),$$

where $F(t) = \int_0^t f(\xi(s))ds$, and there is a standard Wiener process $W(\cdot)$, with respect to which all the other terms of $\hat{\Phi}$ are non-anticipative+ and

$$B(t) = \int_0^t \sigma(\xi(s))dW(s), \quad t < \infty, \text{ w.p.1.}$$

Note. All the jumps of $J(\cdot)$ occur at the $\{\tau_i\}$, $\tau_i = \sum_{j=1}^{i} \delta\tau_j$, $i \geq 1$, but there can be a jump of zero magnitude at any τ_i. For later use (and until mentioned otherwise), we say that $J(\cdot)$ or $\xi(\cdot)$ jump at each τ_i - even if the jump is of zero magnitude. Thus, $J(\cdot)$ jumps at t, if $t = \tau_i$ for some i. Using Skorokhod imbedding, $\xi^h(\tau_i^h) - \xi^h(\tau_i^{h-}) \to \xi(\tau_i) - \xi(\tau_i^-) \equiv \delta\xi(\tau_i)$ w.p.1, as $h \to 0$, each i. Also, to get $W(\cdot)$ we may have to augment the probability space by adding an independent Wiener process.

+By non-anticipative τ_i we mean that the process $I_i(\cdot)$ with values $I_i(t) = I_{\{\tau_i \leq t\}}$ is non-anticipative.

Proof. The continuity, martingale, and representation properties of $B(\cdot)$ are proved as in [3], Theorems 6.3.1 and 6.3.2. In the proof of the latter theorem, we replace $\xi^h(s_i)$ and $\xi(s_i)$ by $(\xi^h(s_i), J^h(s_i), \tau_i^h \cap t)$ and $(\xi(s_i), J(s_i), \tau_i \cap t)$, resp., and use only times s_i at which $\xi(\cdot), J(\cdot)$ are continuous w.p.1. The rest of the details are omitted.

Note that:

$$V^h(x) = E_x\int_0^{\rho_h} k(\xi^h(s))ds + E_x\phi(\xi^h(\rho_h)), \tag{3.5}$$

where $\rho_h = \min\{t: \xi^h(t) \notin G\} = t^h_{N_h}$. The representation (3.5) for $V^h(\cdot)$ in terms of a functional of $\xi^h(\cdot)$ is critical for our method, because we show that the representation (3.5) actually converges to $R(x)$. This also holds for the interpolation of Section 8.

4. Properties of $J(\cdot)$

In this and in the next section, we will give two methods for showing that there is a Poisson measure $\overline{Q}(\cdot,\cdot)$ with the same properties as $Q(\cdot,\cdot)$ (and a corresponding Poisson process $\overline{Q}(\cdot)$) such that $W(\cdot)$ and $\overline{Q}(\cdot)$ are independent and

$$J(t) = \int_0^t\int g(\xi(s^-),\alpha)\overline{Q}(d\alpha\times ds), \quad t < \infty, \text{ w.p.1.}$$

Let $\mathscr{B}_n^h$ denote the σ-algebra determined by $\{\xi_j^h, j \le n, \text{ and } m_j^h \cap n, \text{ all } j\}$. Let $\mathscr{B}^h(t)$ be the σ-algebra determined by $\{\xi^h(s), s \le t, \text{ and } \tau_j^h \cap t, \text{ all } j\}$. Let $\mathscr{B}(t)$, $\mathscr{S}_j^h$ and $\mathscr{S}_j$, resp., be the σ-algebras determined by $\{\Phi(s), s \le t, \text{ and } \tau_j \cap t, \text{ all } j\}$, $\{\xi^h(s), s < \tau_j^h, \text{ and } \tau_k^h, k \le j\}$ and $\{\Phi(s), s < \tau_j, \text{ and } \tau_k, k \le j\}$, resp. Let $N(t,t+\Delta]$ and $N^h(t,t+\Delta]$, resp., denote the number of jumps of $\xi(\cdot)$ and $\xi^h(\cdot)$, resp., in the interval $(t,t+\Delta]$, $\Delta > 0$.

Theorem 4.1. *Assume* (A1.1), (A1.2), *and* (A1.4) *and consider the convergent subsequence of Theorem* 3.2. *Then*

$$
\begin{aligned}
&P_{\mathscr{B}(t)}\{N(t,t+\Delta] = 0\} = \exp - c\Delta \\
(4.1) \qquad &P_{\mathscr{B}(t)}\{N(t,t+\Delta] = 1\} = c\Delta + o(\Delta), \\
&P_{\mathscr{B}(t)}\{N(t,t+\Delta] > 1\} = o(\Delta),
\end{aligned}
$$

where the $o(\Delta)$ *are uniform in* ω,t. *The* $\{\delta\tau_i\}$, $\delta\tau_i = \tau_i - \tau_{i-1}$, *are independent, and each is exponentially distributed with mean* $1/c$. *Also,*

$$
\begin{aligned}
&P_{\mathscr{S}_i}\{\delta\xi(\tau_i) = \xi(\tau_i) - \xi(\tau_i^-) \in A\} = \\
(4.2) \qquad & \\
&p\{\alpha: g(\xi(\tau_i^-),\alpha) \in A\} = \Gamma(\xi(\tau_i^-),A)
\end{aligned}
$$

w.p.1, for each Borel set $A \in R^r$.

<u>Proof</u>. <u>Part 1</u>. We first prove the analog of (4.1) for $\xi^h(\cdot)$. Fix h, and define $I_{i1}, I_{i2}, I_{i3}, I_{i4}$ to be the indicators of the sets $((t,s] = \text{empty if } s \leq t)$.

$$S_{i1} = \{\xi^h(\cdot) \text{ jumps in } (t, t^h_{i-1}]\}$$

$$S_{i2} = \{\xi^h(\cdot) \text{ jumps at } t^h_i\}$$

$$S_{i3} = \{\xi^h(\cdot) \text{ jumps in } [t^h_{i+1}, t+\Delta]\}$$

$$S_{i4} = \{t^h_i \in (t, t+\Delta]\}.$$

Note that S_{i4} is $\mathscr{B}^h_{i-1}$ measurable. Now

$$P_{\mathscr{B}^h(t)}\{N^h(t,t+\Delta] \geq 1\} \leq E_{\mathscr{B}^h(t)} \sum_i I_{i2} I_{i4},$$

$$\text{(4.3)} \qquad E_{\mathscr{B}^h(t)} \sum_i I_{i4} E_{\mathscr{B}^h_{i-1}} I_{i2} = E_{\mathscr{B}^h(t)} \sum_i I_{i4}(1-\exp -c\Delta t^h_{i-1})$$

$$< c\Delta + c \max_y \Delta t^h(y) = c\Delta + \theta(h),$$

by (A1.4). Also

$$P_{\mathscr{B}^h(t)}\{N^h(t,t+\Delta] \geq 2\} = E_{\mathscr{B}^h(t)} \sum_i I_{i4}(1-I_{i1}) I_{i2} I_{i3}$$

$$\text{(4.4)} \qquad \leq E_{\mathscr{B}^h(t)} \sum_i I_{i4}(1-I_{i1}) I_{i2}(c\Delta + \theta(h))$$

$$\leq c^2\Delta^2 + \theta(h),$$

$$(4.5)\qquad P_{\mathscr{B}^h(t)}\{N^h(t,t+\Delta] = 1\} = E_{\mathscr{B}^h(t)} \sum_i (1-I_{i1})I_{i2}(1-I_{i3})I_{i4}.$$

Using $1 - c\Delta + \theta(h) \leq E_{\mathscr{B}_i^h}(1-I_{i3}) \leq 1$, and (which follows from (4.4))

$$E_{\mathscr{B}^h(t)} \sum_i I_{i1}I_{i2}I_{i4} = O(\Delta^2) + \theta(h),$$

we get that the r.h.s. of (4.5) is

$$(4.6)\qquad c\Delta + O(\Delta^2) + \theta(h).$$

By (4.3) and the weak convergence,

$$(4.7)\qquad P\{\xi(\cdot) \text{ jumps at } t\} = 0, \text{ each } t < \infty.$$

Thus, for each t, $\xi(\cdot)$ is continuous at t, w.p.1. Define $I_i(t,t+\Delta;\ x(\cdot))$ to be the function on $D^r[0,\infty) \times (R^+)^\infty = S$ which is 1 if there are i jumps of $x(\cdot)$ on $(t,t+\Delta]$, and is zero otherwise. The generic element of S is $x(\cdot)$, $\{\delta\sigma_i,\ i = 1,2,\ldots\}$ where $\delta\sigma_i$ is the i^{th} interjump interval for a zero or a non-zero jump value. Then, by (4.7), the function is continuous w.p.1 on S, with respect to the $\xi(\cdot)$, $\{\delta\tau_i\}$ measure.

Let $m(\cdot)$ denote a bounded continuous real valued function on the appropriate space, q and $t_1,\ldots,t_q \le t$ arbitrary, and note, by (4.3) - (4.7) and the weak convergence,

$$
\begin{aligned}
&Em(\xi^h(t_j),\ \tau_j^h \cap t,\ j \le q) I_i(t,t+\Delta;\xi^h(\cdot)) \longrightarrow \\
(4.8) \qquad & Em(\xi(t_j),\ \tau_j \cap t,\ j \le q) I_i(t,t+\Delta;\xi(\cdot)),
\end{aligned}
$$

as $h \to 0$.

The I_i in the l.h.s. of (4.8), can be replaced by $(1-c\Delta+O(\Delta^2) + \theta(h))$ when $i = 0$, and by $c\Delta + O(\Delta^2) + \theta(h)$ when $i = 1$.

The same replacements (without the $\theta(h)$) can be made in the r.h.s. of (4.8). Consequently, we have the relation ($i = 0$ or 1)

$$
\begin{aligned}
&Em(\xi(t_j),\tau_j \cap t,\ j \le q) I_i(t,t+\Delta;\xi(\cdot)) \\
&\quad = Em(\xi(t_j),\ \tau_j \cap t,\ j \le q) P_{\mathscr{B}(t)}\{N(t,t+\Delta] = i\} \\
&\quad = Em(\xi(t_j),\ \tau_j \cap t,\ j \le q)[(1-c\Delta+o(\Delta)) \quad \text{or} \quad (c\Delta+o(\Delta))],
\end{aligned}
$$

which, together with the arbitrariness of $m(\cdot)$, q, and t_i, $i \le q$, imply that (4.1) holds with r.h. sides, $1 - c\Delta + o(\Delta)$, $c\Delta + o(\Delta)$, $o(\Delta)$, resp. From this, it is easy to show that (4.1) holds as stated. The sentence following (4.1) follows from (4.1).

<u>Part 2</u>. Fix i and let A be a closed Borel set in R^r such that it is the closure of its interior and

$$\begin{aligned} &Ep\{\alpha: g(\xi(\tau_i^-),\alpha) \in \partial A\} \\ &= P\{\delta\xi(\tau_i) \in \partial A\} = 0. \end{aligned}$$

(I.e., if q has the distribution $p(\cdot)$ and is independent of $\xi(\tau_i^-)$, then $P\{g(\xi(\tau_i^-),q) \in \partial A\} = 0.$) Then

$$p\{\alpha: g(\xi(\tau_i^-),\alpha) \in \partial A\} = 0 \text{ for almost all } \xi(\tau_i^-) \text{ values, } (\xi(\cdot), \tau_i \text{ measure}). \tag{4.9}$$

By (4.9), and the convergence $\xi^h(\tau_i^h) \to \xi(\tau_i^-)$ w.p.1 (Skorokhod imbedding used), and the properties of A and ∂A,

$$p(\alpha: g(\xi^h(\tau_i^{h-}),\alpha) \in A) \to p(\alpha: g(\xi(\tau_i^-),\alpha) \in A) \tag{4.10}$$

for almost all $\xi(\tau_i^-)$ values $(\xi(\cdot),\tau_i$ measure). Now, note the following $(m(\cdot),q$, are as in Part 1, and the t_j are $< \infty$):

$$\begin{aligned} &Em(\xi^h(t_j \cap \tau_i^{h-}); \ \tau_j^h \cap \tau_i^h, \ j \le q) I_{\{\delta\xi_i^h \varepsilon A\}} \\ &= Em(\xi^h(t_j \cap \tau_i^{h-}); \ \tau_j^h \cap \tau_i^h, \ j \le q) P_{\mathcal{S}_i^h}\{\delta\xi_i^h \in A\} \end{aligned} \tag{4.11}$$

The l.h.s. of (4.11) converges to

$$\begin{aligned} &Em(\xi(t_j \cap \tau_i^-),\ \tau_j \cap \tau_i,\ j \le q) I_{\{\delta\xi(\tau_i)\varepsilon A\}} = \\ (4.12)\qquad &Em(\xi(t_j \cap \tau_i^-),\ \tau_j \cap \tau_i,\ j \le q) P_{\mathcal{S}_i}\{\delta\xi(\tau_i)\ \varepsilon\ A\}. \end{aligned}$$

Also

$$(4.13)\quad P_{\mathcal{S}_i^h}\{\delta\xi_i^h\ \varepsilon\ A\} \to p(\alpha: g(\xi(\tau_i^-),\alpha)\ \varepsilon\ A),$$

(a $\mathcal{S}_i$ measurable function) for almost all $\xi(\tau_i^-)$ values ($\xi(\cdot)$, τ_i measure). Thus, the l.h.s. of (4.11) also converges to (4.12) with the r.h.s. of (4.13) replacing the indicator in (4.12). Now, the arbitrariness of $m(\cdot)$, q, t_j, $j \le q$, implies (4.2) for A of the chosen class. Since such A generate the Borel algebra over R^r, (4.2) holds as stated. Q.E.D.

<u>Theorem 4.2</u>. <u>Under</u> (A1.1) <u>to</u> (A1.4), $\xi(\cdot)$ <u>and</u> $X(\cdot)$ <u>(given by</u> (1.1)) <u>induce the same measure on</u> $D^r[0,\infty)$.

<u>Proof</u>. Define a jump time of $X(\cdot)$ to be a jump time of $Q(\cdot)$. Then the interjump times for $X(\cdot)$ are mutually independent, and exponentially distributed with mean value $1/c$.

Furthermore, the conditional distribution of the value of the i^{th} jump of $X(\cdot)$ (given $X(s)$, $s < i^{th}$ jump time, and

j^{th} jump time, $j \le i$) is the same as for $\xi(\cdot)$. Between jumps the processes evolve as diffusions. Thus, by (A1.3), the induced measures are the same for $\xi(\cdot)$ and $X(\cdot)$. In particular, the $\xi(\cdot)$ measure does not depend on the subsequence. Q.E.D.

5. An Alternative Representation for $J^h(\cdot)$ and $\xi^h(\cdot)$

The representation to be developed in this section is particularly useful when we treat the control case, for it will allow us to prove optimality theorems along the lines of those in [3], Chapters 8 and 9.

Theorem 5.1. *Assume* (A1.1), (A1.2) *and* (A1.4). *Let* $\hat{Q}^h(\cdot)$ *and* $\hat{Q}^h(\cdot,\cdot)$ *denote a Poisson process (and corresponding measure) satisfying* (A1.2). *Then we can write*

$$\xi_n^h = x + F_n^h + B_n^h + \hat{J}_n^h + \hat{E}_n^h , \tag{5.1}$$

where $\hat{E}^h(\cdot)$, *the interpolation of* $\{\hat{E}_n^h\}$, *converges weakly to the zero process and*

$$\hat{J}_n^h = \sum_{i=0}^{n-1} \hat{j}_i^h,$$

$$\hat{j}_i^h = \int_{t_i^h}^{t_{i+1}^h} \int g(\xi^h(t_i^h),\alpha)\hat{Q}^h(d\alpha\times ds) = \tag{5.2}$$

$$= \int_{t_i^h}^{t_{i+1}^h} \int g(\xi^h(s^-),\alpha)\hat{Q}^h(d\alpha\times ds),$$

and $\hat{J}^h(\cdot)$ is tight on $D^r[0,\infty)$. (The integrations $\int_{t_i^h}^{t_{i+1}^h}$ are always over the interval $(t_i^h, t_{i+1}^h]$. Only the jumps of $\hat{Q}(\cdot)$ on $(t_i^h, t_{i+1}^h]$ play a role. Define $\xi^h(0^-) = \xi^h(0)$, etc.)

Proof. Define $s_i^h = \min\{t: t > t_i^h, \hat{Q}^h(t) - \hat{Q}^h(t^-) \neq 0\}$, and

$$\bar{\varepsilon}_i^h = -\int_{s_i^h \cap t_{i+1}^h}^{t_{i+1}^h} \int g(\xi_i^h, \alpha)\hat{Q}^h(d\alpha \times ds).$$

Then, for any Borel set A,

$$P\{\hat{j}_i^h + \bar{\varepsilon}_i^h \in A | \xi_j^h, j \le i, s_i^h \le t_{i+1}^h\} = \Gamma(\xi_i^h, A), \tag{5.3a}$$

and

$$P\{\hat{j}_i^h + \bar{\varepsilon}_i^h \in b^h(j_1,\dots,j_r) | \xi_j^h, j \le i, s_i^h \le t_{i+1}^h\} = \tag{5.3b}$$

$$= \Gamma^h(\xi_i^h, j_1,\dots,j_r).$$

Let us construct a new chain, also called $\{\xi_n^h\}$, as follows. If $\hat{Q}^h(\cdot)$ does not jump in $(t_i^h, t_{i+1}^h]$, then let ξ_{i+1}^h evolve from ξ_i^h using the Markov law $p^h(\cdot,\cdot)$, as before. If $\hat{Q}^h(\cdot)$ does jump in $(t_i^h, t_{i+1}^h]$, then set

$$\xi_{i+1}^h - \xi_i^h = (j_1'h,\ldots,j_r'h) \quad \text{if} \quad \hat{j}_i^h + \bar{\varepsilon}_i^h \in b^h(\xi_i^h, j_1,\ldots,j_r).$$

The law of the new chain is exactly the same as the law of the former chain.

The difference $\tilde{\varepsilon}_i^h = (\xi_{i+1}^h - \xi_i^h) I_{\{\text{jump at } i+1\}} - (\hat{j}_i^h + \bar{\varepsilon}_i^h)$ is due to the method of discretizing the jumps, and $\{\tilde{\varepsilon}^h(\cdot)\}$ converges weakly to the zero process, as $h \to 0$. Define $\hat{\varepsilon}_i^h = -\varepsilon_i^h + \bar{\varepsilon}_i^h + \tilde{\varepsilon}_i^h$, with interpolation $\hat{E}^h(\cdot)$. Then (5.1) holds, and $\hat{E}^h(\cdot) \to$ zero process weakly, as $h \to 0$. Tightness of $\{\hat{J}^h(\cdot)\}$ follows from the properties of $\{\hat{Q}^h(\cdot)\}$. Q.E.D.

Theorem 5.2. Assume the conditions of Theorem 5.1, and let+ $g(\cdot,\cdot)$ be continuous. Then $\{\xi^h(\cdot), B^h(\cdot), \hat{J}^h(\cdot), \hat{E}^h(\cdot), \hat{Q}^h(\cdot)\}$ is tight on $D^{4r+r'}[0,\infty)$. Let h index a convergent subsequence, with limit $\xi(\cdot), B(\cdot), J(\cdot), 0, Q(\cdot)$. Then $\hat{Q}(\cdot)$ is a Poisson process of the type of (A1.2) (with corresponding Poisson measure $\hat{Q}(\cdot,\cdot)$), and there is a standard Wiener process $W(\cdot)$ (we may possibly have to augment the space, by adding an independent Wiener process) such that $\hat{Q}(\cdot)$ and $W(\cdot)$ are independent, $\xi(\cdot), B(\cdot), J(\cdot)$ are non-anticipative with respect to $\hat{Q}(\cdot)$ and $W(\cdot)$, and

$$\xi(t) = x + \int_0^t f(\xi(s))ds + B(t) + J(t), \tag{5.4}$$

$$B(t) = \int_0^t \sigma(\xi(s))dW(s) \tag{5.5}$$

$$J(t) = \int_0^t \int g(\xi(s^-),\alpha)\hat{Q}(d\alpha \times ds). \tag{5.6}$$

+The continuity has and (also in Theorems 6.1 and 8.2) can be removed by use of a more "careful" choice of the discretization for the jump.

<u>Proof</u>. The tightness follows from Theorems 3.1 and 5.1. Also, $\hat{Q}(\cdot)$ is a Poisson process of the asserted type since each $\hat{Q}^h(\cdot)$ is. The representation (5.4) follows from the weak convergence.

A modification of the proof of Theorem 6.3.2 in [3] yields that $B(\cdot)$ is a continuous martingale with respect to $\{\hat{\mathscr{B}}(t)\}$, where $\hat{\mathscr{B}}(t)$ is the smallest σ-algebra which measures $\xi(s)$, $J(s)$, $\hat{Q}(s)$, $s \le t$. The quadratic covariation of $B(\cdot)$ is $\int_0^t 2a(\xi(s))ds$. Augment the probability space by adding $\psi(\cdot)$, an independent Wiener process, and let $\bar{\mathscr{B}}(t)$ be the smallest σ-algebra which measures $\xi(s)$, $J(s)$, $\hat{Q}(s), \psi((s)$, $s \le t$. By the method of Chapter 1.4.4 and Theorem 6.3.2 of [3], we obtain a $W(\cdot)$ (non-anticipatively) from $\xi(\cdot)$, $B(\cdot)$ such that $\{(W(t), \bar{\mathscr{B}}(t)\}$ is a Wiener process and (5.5) holds for all t, w.p.1. This, together with the fact that $\hat{Q}^h(\cdot)$ is a Poisson process, implies that $\{\hat{Q}(t+s) - \hat{Q}(t), W(t+s) - W(t), s \ge 0\}$ is independent of $\{\hat{Q}(u), W(u), u \le t\}$ for each t. Thus $\{W(\cdot), \hat{Q}(\cdot)\}$ has independent increments (see also Section 9) and since $W(\cdot)$ is continuous and $\hat{Q}(\cdot)$ is a pure jump process, by Gikhman and Skorokhod [5], p. 271, they are mutually independent.

We have only to prove the representation (5.6). Define $\xi(0^-) = \xi(0)$, and

$$A_n^{\Delta} = \sum_{i=0}^{n-1} \int_{i\Delta}^{i\Delta+\Delta} [g(\xi(i\Delta),\alpha) - g(\xi(s^-),\alpha)]\hat{Q}(d\alpha \times ds)$$

$$A_n^{h,\Delta} = \sum_{i=0}^{n-1} \int_{i\Delta}^{i\Delta+\Delta} [g(\xi^h(i\Delta),\alpha) - g(\xi^h(s^-),\alpha)]\hat{Q}^h(d\alpha \times ds),$$

and let $A(\Delta,\cdot)$, $A^h(\Delta,\cdot)$ denote the functions whose values are given by the A_n^{Δ} and $A_n^{h,\Delta}$, on the time intervals $[n\Delta, n\Delta+\Delta)$, resp. We can show that (using Skorokhod imbedding)

$$\lim_{\Delta\to 0} A(\Delta,\cdot) = \text{zero process w.p.1}$$

$$\lim_{\Delta\to 0} \lim_{h\to 0} A^h(\Delta,\cdot) = \text{zero process w.p.1}$$

Thus, to prove (5.6), we only need to show that

$$B(\Delta,h,i) \equiv \int_{i\Delta}^{i\Delta+\Delta} g(\xi^h(i\Delta),\alpha)\hat{Q}^h(d\alpha \times ds)$$

$$\to \int_{i\Delta}^{i\Delta+\Delta} g(\xi(i\Delta),\alpha)\hat{Q}(d\alpha \times ds) \equiv B(\Delta,i), \text{ w.p.1},$$

for each Δ, i (Skorokhod imbedding assumed, as usual, where convenient).

Now

$$B(\Delta,h,i) = \sum_{j\varepsilon\Delta} g(\xi^h(i\Delta),q_j^h)$$

$$B(\Delta,i) = \sum_{j\varepsilon\Delta} g(\xi(i\Delta),q_j),$$

where q_j^h (resp. q_j) are the j^{th} jumps of $\hat{Q}^h(\cdot)$, and $\hat{Q}(\cdot)$, resp., and, by $j \in \Delta$, we mean that we sum over only the j for which the j^{th} jump occurs in the interval $(i\Delta, i\Delta+\Delta]$.

By weak convergence, we have $q_j^h \to q_j, \xi^h(i\Delta) \to \xi(i\Delta)$ (w.p.1), and using the fact that

$P\{\text{jump of } \hat{Q}(\cdot) \text{ at } i\Delta \text{ or } i\Delta + \Delta\} = 0$, and the continuity of $g(\cdot,\cdot)$ we get that

$$B(\Delta,h,i) \to B(\Delta,i) \quad \text{w.p.1, as } h \to 0. \qquad \text{Q.E.D.}$$

<u>Remark</u> on a useful representation for $\{\beta_n^h\}$. Augment the space on which $\{\xi_n^h\}$ are defined by adding an independent standard R^r valued Wiener process $\psi(\cdot)$. In Theorem 6.6.1 of [3] (where $g(\cdot,\cdot) = 0$), it was shown that we can write (for some $\{\varepsilon_n^h\}$)

$$\beta_n^h = \sigma(\xi_n^h)\delta W_n^h + \varepsilon_n^h,$$

where the interpolation of ε_n^h converges weakly to the zero process and δW_n^h (and ε_n^h) depend on $\xi_n^h, \beta_n^h, \psi(t_{n+1}^h) - \psi(t_n^h)$. Also, the process $W^h(\cdot)$, defined by the interpolation of $W_n^h = \sum_{i=0}^{n-1} \delta W_n^h$, converges weakly to a standard R^r valued Wiener process. $W(\cdot)$, and (5.5) holds with this $W(\cdot)$. We can do the same thing here, defining δW_n^h <u>exactly</u> as it was defined in [3], Theorem 6.6.1, and Theorem 5.2 holds if $W(\cdot)$ is constructed as a limit of $\{W^h(\cdot)\}$. We then can write

$$\text{(5.7)} \qquad \xi_{n+1}^h = \xi_n^h + f(\xi_n^h)\Delta t_n^h + \sigma(\xi_n^h)\delta W_n^h + \int\int_{t_n^h}^{t_{n+1}^h} g(\xi_n^h,\alpha)\hat{Q}^h(d\alpha \times ds) + E_n^h,$$

for some $\{E_n^h\}$ such that $E^h(\cdot) \to$ zero process weakly, as $h \to 0$.

Remarks on an Alternative Chain for that of Section 2.

It is conceivable that the coefficients of all terms in (2.2) such that $\sum_i e_i j_i' = 0$, are not zero. Then, when there is a jump, the jump could be strictly zero. If these probabilities are small, then its' not important - but, if not, then we may be able to save some computation by using a slightly different chain $\{\xi_n^h\}$. Define

$$\bar{c}(x) = cp\{\alpha: g(x,\alpha) \neq 0\}$$

$$\bar{\Gamma}(x,A) = p\{\alpha: g(x,\alpha) \in A - \{0\}\}/p\{\alpha: g(x,\alpha) \neq 0\}.$$

Assume that $\bar{c}(\cdot)$ is continuous. Equation (1.3) can be written as

$$\mathscr{L}V(x) + \bar{c}(x) \int [V(x+\alpha) - V(x)]\bar{\Gamma}(x,d\alpha) + k(x) = 0. \tag{5.8}$$

Now, proceeding exactly as was done in Section 2, and defining

$$\Delta t^h(x) = h^2/[Q_h(x) + \bar{c}(x)h^2]$$

$$p^h(x) = 1 - \exp - \bar{c}(x)\Delta t^h(x),$$

yields (2.4), where all other terms are as defined there, except that $\bar{\Gamma}^h$ (defined as Γ^h was defined, but using $\bar{\Gamma}$ in lieu of Γ) replaces Γ^h there.

The $\Delta t^h(x)$ are larger than before - since we have eliminated the zero jumps. With the new definitions, Theorems 3.1 and 3.2 remain valid. Theorem 4.1 needs to be modified as follows. Here, let $\sigma_1, \ldots,$ denote the times of the actual (not zero) jumps of $\xi(\cdot)$, and let $\mathscr{D}_i$ be the smallest σ-algebra which measures $\xi(s)$, $s < \sigma_i$, σ_j, $j \leq i$. The second line of (4.1) remains valid, if c is replaced by $\bar{c}(\xi(t))$, and the third line remains valid. The right side of the first line of (4.1) must be replaced by

$$(5.9) \qquad E_{\xi(t)} \exp - \int_t^{t+\Delta} \bar{c}(\bar{\xi}(s))ds,$$

where $\bar{\xi}(\cdot)$ is a diffusion process $d\bar{\xi} = f(\bar{\xi})dt + \sigma(\bar{\xi})dw$ on $[t,\infty)$, with initial condition $\xi(t)$, but which is otherwise independent of $\xi(\cdot)$.

Equation (4.2) must be replaced by (what is expected)

$$(5.10) \qquad P_{\mathscr{D}_i}\{\xi(\sigma_i) - \xi(\sigma_i^-) \in A\} = \bar{\Gamma}(\xi(\sigma_i^-),A)$$

for each Borel A.

Theorem 4.2 also remains valid for the following reason. Between jumps, both $X(\cdot)$ and $\xi(\cdot)$ behave as diffusions, and the conditional distribution of the jumps is the same, where we now define a jump to be an actual non-zero jump (of $\xi(\cdot)$ or of $X(\cdot)$). Furthermore, the distribution of the interjump times is the same for both, namely

$$(5.11)\qquad P\{\sigma_{i+1}-\sigma_i > t \mid \xi(s),\ s \le \sigma_i,\ \xi(\sigma_i) = y\}$$
$$= E_y \exp - \int_0^t \overline{c}(\overline{\xi}(s))ds.$$

which is also the corresponding distribution for $X(\cdot)$.

The details are similar to those of the foregoing proofs, except for (5.9) and (5.11), which are a little more involved. Also, the construction of (5.7) can be carried out for the new chain, and Theorems 5.1 and 5.2 remains valid.

6. The Case $c = \infty$

We now consider the case where the jump rate is infinite, but where "most" jumps are very small.

Assume

(A6.1) $\pi(A) < \infty$ *for each* A *which is disjoint from the origin*. (*Then, for each* $\varepsilon > 0$, $\pi(\cdot)$ *is a finite measure on* $R^r - \{x: |x| \le \varepsilon\}$.)

(A6.2) There is a real K such that $|g(x,\alpha)| \le K|\alpha|$.

(A6.3)
$$\int |\alpha|\pi(d\alpha) < \infty$$
$$\int_{|\alpha|\le\varepsilon} |\alpha|\pi(d\alpha) \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Other sets of assumptions, besides A6.1-A6.3 can be used. For example, each component of the vector $Q(\cdot)$ can be treated separately with the appropriate assumptions on $g(\cdot,\cdot)$.

For each $\varepsilon > 0$, let $g_\varepsilon(\cdot,\cdot)$ denote a function satisfying (A1.1) and (A6.2) (K independent of ε), and which is zero when $|\alpha| \leq \varepsilon/2$, and equal to $g(\cdot,\cdot)$ for $|\alpha| \geq \varepsilon$. If $\hat{\Gamma}(x,R^r) = \infty$ for some x, then we would have $\Delta t^h(x) = 0$, and the approximation procedure in Section 2 has to be modified. (This makes intuitive sense, for the interpolation interval should decrease as the jump rate increases.) Choose some sequence (to be further restricted below) $\varepsilon_h \to 0$, as $h \to 0$, and define $\hat{\Gamma}_\varepsilon(x,A) = \pi(\alpha\colon g_\varepsilon(x,\alpha) \in A)$. With difference interval h, we discretize (6.1) in lieu of the current form of (1.3).

$$\hat{\Gamma}_{\varepsilon_h}(x,A) = \pi(\alpha\colon g_{\varepsilon_h}(x,\alpha) \in A)$$

$$\mathscr{L}V(x) + \int [V(x+\alpha) - V(x)]\hat{\Gamma}_{\varepsilon_h}(x,d\alpha) + k(x) = 0, \quad x \in G. \tag{6.1}$$

Define $\hat{\Gamma}^h_{\varepsilon_h}(x,j_1,\ldots,j_r)$ as $\Gamma^h(x,j_1,\ldots,j_r)$ was defined, but using $\hat{\Gamma}_{\varepsilon_h}$ in lieu of Γ. Define $c_h(x) = \sum_{\{j_i\}} \hat{\Gamma}^h_{\varepsilon_h}(x,j_1,\ldots,j_r)$, where the sum is over all $(j_1,\ldots,j_r)$ for which $(j_1',\ldots,j_r') \neq 0$. The jumps of zero value are to be deleted, and we will use

$$\Delta t^h(x) = h^2/[Q_h(x) + h^2 c_h(x)]$$

$$p^h(x) = [1 - \exp - c_h(x)\Delta t^h(x)] \quad \text{or} \quad [c_h(x)\Delta t^h(x)],$$

analogously to (2.3).

Define $p^h(x,y)$ as in Section 2. Then we get the discretized form (2.4), but with $\hat{\Gamma}^h_{\varepsilon_h}(x,j_1,\ldots,j_r)/c_h(x)$, $c_h(x)$ and the new $\Delta t^h(x)$ and $P^h(x)$ replacing $\Gamma^h(x,j_1,\ldots,j_r)$, c and the former $\Delta t^h(x)$ and $P^h(x)$.

Theorem 6.1. Choose ε_h such that

$$\max_x c_h(x)\Delta t^h(x) \to 0, \max_x \pi(\alpha: |\alpha| \geq \varepsilon_h/2)\Delta t^h(x) \to 0, \text{ as } h \to 0.$$

Assume[+] (A1.1) to (A1.4) and (A6.1) to (A6.3). Then (omitting the assertions concerning $\delta\tau_i^h$ and $\delta\tau_i$) Theorems 3.1, 3.2, 5.1 and 5.2 (if $g(\cdot,\cdot)$ is continuous) continue to hold: in Theorems 5.1, 5.2, we need to add to (the previously defined) $\hat{E}^h(\cdot)$ a process $\mathring{E}^h(\cdot)$, which also tends to zero weakly as $h \to 0$. Also, the remarks after Theorem 5.2, concerning $W^h(\cdot)$ and its limit $W(\cdot)$, continue to hold.

Proof. Tightness of $\{F^h(\cdot),B^h(\cdot)\}$ follows from Theorem 3.1. The value of ε_h is chosen to assure that $E^h(\cdot)$ (see Theorem 3.1) converges to the zero process. Let $J^{h,\delta}(\cdot)$ denote $J^h(\cdot)$, but where all jumps of absolute magnitude $\leq \delta$ are deleted. Then by Theorem 3.1, for each $\delta > 0$, $\{J^{h,\delta}(\cdot)\}$ is tight. By (A6.2) - (A6.3), there is a real K_1 such that for each $T < \infty$, $\delta > 0$,

$$E \max_{t \leq T} |J^h(t) - J^{h,\delta}(t)| \leq K_1 \delta T.$$

[+]Where A1.2 is modified to take account of the jump properties of $\hat{Q}(\cdot)$ assumed in this Section.

The last two sentences imply that, for any subsequence of $\{J^h(\cdot)\}$ we can find a further subsequence that converges weakly in $D^r[0,\infty)$. Since tightness is used only to get weak convergence, we can assume that $\{J^h(\cdot)\}$ is tight. Theorems 3.1 and 3.2 (except for the assertions concerning $\delta\tau_i^h, \delta\tau_i$) follow from these remarks.

For Theorems 5.1 and 5.2 (with the appropriate generalization of (A1.2)) define $\bar{\varepsilon}_i^h$ using $g_{\varepsilon_h}(\cdot,\cdot)$ in lieu of $g(\cdot,\cdot)$, and define $s_i^h = \min\{t: t > t_i^h, |\hat{Q}^h(t) - \hat{Q}^h(t^-)| \geq \varepsilon_h/2\}$. Define

$$\mathring{\varepsilon}_i^h = \int_{t_i^h}^{t_{i+1}^h} [g_{\varepsilon_h}(\xi_i^h,\alpha) - g(\xi_i^h,\alpha)]\hat{Q}^h(d\alpha \times ds),$$

$$\mathring{E}_n^h = \sum_{j=0}^{n-1} \varepsilon_j^h .$$

We continue to define $\hat{j}_i^h$ using $g(\cdot,\cdot)$, as in (5.2). Both $\mathring{E}^h(\cdot)$ and $\bar{E}^h(\cdot) \to$ zero process as $h \to 0$, by (A6.1) to (A6.3). Modify $\hat{E}_n^h$ in (5.1) by adding $\mathring{E}_n^h$. Then Theorem 5.1 continues to hold.

For Theorem 5.2, the assertions there concerning $A(\Delta,\cdot)$ and $A^h(\Delta,\cdot)$ continue to hold in the present case, and we only outline the convergence of $B(\Delta,h,i)$ to $B(\Delta,i)$. By (A6.1) to (A6.3), both

$$E \sum_{\substack{j\varepsilon\Delta \\ |q_j^h|\le\delta}} |g(\xi^h(i\Delta),q_j^h)|, \quad E \sum_{\substack{j\varepsilon\Delta \\ |q_j|\le\delta}} |g(\xi(i\Delta),q_j)|$$

tend to zero as $\delta \to 0$, uniformly in h. So we only need to show convergence when $|q_j|$ and $|q_j^h|$ are restricted to be $\geq \delta$, for δ arbitrarily close to zero. Choose any $\delta > 0$ such that $\pi(\alpha: |\alpha| = \delta) = 0$. Then, by a weak convergence argument (using Skorokhod imbedding), as in Theorem 5.2, we get

$$\sum_{\substack{j\varepsilon\Delta \\ |q_j^h|>\delta}} g(\xi^h(i\Delta),q_j^h) \to \sum_{\substack{j\varepsilon\Delta \\ |q_j|>\delta}} g(\xi(i\Delta),q_j)$$

w.p.1, as $h \to 0$. Q.E.D.

<u>Remark</u>. The fact that Theorem 5.2 holds implies that $\xi(\cdot)$ and $X(\cdot)$ have the same distributions under (A1.3), if $g(\cdot,\cdot)$ is continuous. Suppose that $\xi^h(\cdot)$ is constructed by the discretization of this section, but without assuming continuity of $g(\cdot,\cdot)$ in the second argument. Then Theorem 4.2 still holds. The proof is more complicated and is omitted.

7. Convergence of the Functionals $V^h(x)$ to $R(x)$

The chain $\{\xi_n^h\}$ can be used to approximate a large class of functionals of $X(\cdot)$, and of solutions to equations such as (1.3).

Theorem 7.1 (see [3], Theorem 6.4.1 and [6], Chapters 2 and 4). Under (A1.1) to (A1.4) (and (A6.1) - (A6.3), if $c = \infty$), if $F(\cdot)$ is any real valued bounded and measurable function on $D^r[0,\infty)$ which is a.e. continuous with respect to the measure induced by $\xi(\cdot)$ or $X(\cdot)$, with initial condition x, then

$$E_x F(\xi^h(\cdot)) \to E_x F(\xi(\cdot)), \text{ as } h \to 0.$$

The theorem holds if uniform integrability of $\{F(\xi^h(\cdot))\}$ replaces boundedness, and the convergence is uniform on compact x sets.

To treat $V^h_{\cdot}(x) \to R(x)$, in particular, some more conditions are needed, since we need to know that our functionals are a.e. continuous (as in the theorem) and that the approximations are uniformly integrable. Define $\tau(\cdot): D^r[0,\infty) \to [0,\infty]$ by $\tau(x(\cdot)) = \inf\{t: x(t) \not\in G\}$. We need to assume that

(A7.1) $\tau(\cdot)$ is continuous w.p.1 ($X(\cdot)$ measure, $X(0) = x$); i.e., that, w.p.1, there are no path tangencies at the point of contact of $X(\cdot)$ with ∂G.

We also need that

(A7.2) $E_x \rho_h$ <u>is uniformly bounded in</u> h (<u>to get uniform integrability</u>). <u>This is implied by</u>

$$\inf_{x \varepsilon \bar{G}} P_x\{\text{escape time of } X(\cdot) \text{ from } \bar{G} \leq T\} > 0$$

<u>for some</u> $T < \infty$. <u>See</u> [3], Chapter 6.4.

Now, under conditions (A1.1) to (A1.4) (and (A6.1) to (A6.3), and continuity[+] of $g(\cdot,\cdot)$, if $c = \infty$), and (A7.1) - (A7.2), and $\xi^h(\cdot) = x$, we have that $\xi^h(\cdot) \to \xi(\cdot)$ (which has the law of $X(\cdot)$, $X(0) = x$) as $h \to \infty$, and

$$\phi(\xi^h(\rho_h)) \to \phi(\xi(\tau))$$

$$\int_0^{\rho_h} k(\xi^h(s))ds \to \int_0^{\tau} k(\xi(s))ds$$

in distribution (and the mean values also converge; i.e., $V^h(x) \to R(x)$), as $h \to 0$.

8. An Alternative Interpolation $\xi^h(\cdot)$

In [3] and in Section 2, the interpolation $\xi^h(\cdot)$ is constant on time intervals $\{\Delta t_i^h\}$. Given $\xi_i^h = x$, the interpolation interval $[t_i^h, t_{i+1}^h)$ is <u>known</u>, and $\xi^h(\cdot)$ is not a

---

[+]Theorem 6.1 assumed the continuity, but Theorem 4.2 still holds when $c = \infty$, although the proof was not given. Under (the extended) Theorem 4.2, the continuity of $g(\cdot,\cdot)$ can be dropped, in favor of (A1.1).

Markov process. For some purposes, it would be convenient if $\xi^h(\cdot)$ were Markovian. In this section, a right continuous Markov interpolation, with random (exponentially distributed) interpolation intervals, will be developed. It has the additional advantage that $\sup_x \Delta t^h(x)$ need not be finite, and the condition (A1.4) can be dropped. We assume $c < \infty$, although there is an analogous development for the case of Section 6. Unless otherwise specified, symbols retain their earlier definitions. First, the simple case $g(\cdot,\cdot) \equiv 0$ will be treated.

8.1. <u>Case I</u> $g(\cdot,\cdot) = Q(\cdot) \equiv 0$ (the case of [3], Chapter 6). Let $\Delta t^h(x) = h^2/Q_h(x)$. Define $\xi^h(\cdot)$ (we use the same symbol for the new interpolation) to be a <u>Markov jump</u> process as follows. Let $\mathscr{B}^h(t)$ be the smallest σ-algebra which measures $\xi^h(s)$, $s \le t$, set $\tau_0^h = 0$ and, for $i \ge 1$, let τ_i^h and $\tau_i^h - \tau_{i-1}^h = \delta\tau_i^h$ denote the i^{th} jump time and interjump time, resp., of $\xi^h(\cdot)$. Assume that

$$(8.1)\quad \begin{aligned} &P\{\xi^h(\cdot) \text{ changes in } (t,t+\Delta] \,|\, \mathscr{B}^h(t), \xi^h(t) = x\} = 1 - \exp - (\Delta/\Delta t^h(x)), \\ &P\{\xi^h(\tau_i^h) - \xi^h(\tau_{i-1}^h) = y \,|\, \xi^h(\tau_{i-1}^h) = x\} = p^h(x,y), \quad \text{all } x,y. \end{aligned}$$

By (8.1),

$$(8.2)\quad E[\tau_{i+1}^h - \tau_i^h \,|\, \xi^h(s),\ s \le \tau_i^h] = \Delta t^h(\xi^h(\tau_i^h)).$$

Owing to (8.2), and to the law of $\xi^h(\tau_i^h) - \xi^h(\tau_{i-1}^h)$, <u>the solution</u>

to (2.4) with ($P^h(\cdot) = c = \Gamma^h = 0$ here) still has the representation (3.5), where ρ_h is the escape time of the new $\xi^h(\cdot)$ from G. The fact that $V^h(x)$ still has the representation (3.5) is critical for our method. This, together with the (to be proved under (A1.3)) fact that the sequence of new "random" interpolations converges weakly to (1.1), implies that the results of Section 7 still hold for our new interpolation. This justifies the use of this new interpolation also. Thus: cost functional for chain = cost functional for interpolation, which, in turn, converges to $R(x)$. The same is true for the case $g, Q \not\equiv 0$, discussed below. The actual computations of the approximations to functionals of $X(\cdot)$, or to solutions of (1.3) will still be done using the law of the chain. The interpolation is used only for the theoretical arguments in the convergence proofs.

Define $\bar{\Sigma}_h(x) = \sigma(x)\sigma'(x) + h\tilde{f}(x) = \Sigma_h(x) - f(x)f'(x)(\Delta t^h(x))^2$ $= \Sigma_h(x) - E[\xi_{n+1}^h - \xi_n^h | \xi_n^h = x]\cdot(E[\xi_{n+1}^h - \xi_n^h | \xi_n^h = x])'$. Owing to the nature of the interpolation $\xi^h(\cdot)$, the $(\text{mean})^2$ term does not appear in the expression for the quadratic covariation.

Theorem 8.1. Assume (A1.1), (A1.2), and fix $\xi_0^h = \xi^h(0) = x$. Let $g(\cdot,\cdot) = 0$. Then we can write

$$\xi^h(t) = x + F^h(t) + B^h(t), \tag{8.3}$$

where

$$F^h(t) = \int_0^t f(\xi^h(s))ds$$

$$B^h(t) = \int_0^t [d\xi^h(s) - f(\xi^h(s))ds] \equiv \int_0^t dB^h(s),$$

and $B^h(\cdot)$ is a martingale with quadratic covariation

$$\int_0^t \bar{\Sigma}_h(\xi^h(s))ds.$$

The sequence $\{\xi^h(\cdot), F^h(\cdot), B^h(\cdot)\}$ is tight on $D^{3r}[0,\infty)$, and for any weakly convergent subsequence, with limit denoted by $\xi(\cdot)$, $F(\cdot)$, $B(\cdot)$, there is a Wiener process $W(\cdot)$ such that $\xi(\cdot)$, $B(\cdot)$ are non-anticipative with respect to $W(\cdot)$ and

$$\xi(t) = x + F(t) + B(t) = x + \int_0^t f(\xi(s))ds + \int_0^t \sigma(\xi(s))dW(s). \tag{8.4}$$

There is a martingale process $W^h(\cdot)$ with quadratic covariation It such that

$$B^h(t) = \int_0^t \sigma(\xi^h(t))dW^h(s) + E^h(t) + \int_0^t [\bar{\sigma}_h(\xi^h(s)) - \sigma(\xi^h(s))]dW^h(s), \tag{8.5}$$

where the last two processes in (8.5) tend to the zero process

weakly, as $h \to 0$, and $\bar\sigma_h(x)\bar\sigma_h'(x) = \bar\Sigma_h(s)$, $\bar\sigma_h(x) \to \sigma(x)$ as $h \to 0$. The sequence $\{W^h(\cdot)\}$ is tight. Also, any weak limit, $W(\cdot)$, is a Wiener process with the properties of the $W(\cdot)$ above. Under (A1.3), the limit $\xi(\cdot)$ is the unique solution to (1.1) (in the sense of distributions).

Proof. The proof is close to that given for the original interpolation in [3], Chapter 6.3 and 6.6, and only an outline will be given.

Part 1. Equation (8.3) holds by the definitions. Fix h. Note that

$$E[\xi^h(t+\Delta) - \xi^h(t) | \mathscr{B}^h(t), \xi^h(t) = x] = (f(x)\Delta t^h(x))\frac{\Delta}{\Delta t^h(x)} + o(\Delta)$$

(= $o(\Delta)$ + conditional (on a jump) average change in $\xi^h(\cdot)$ over the interval $(t,t+\Delta]$ times conditional probability of jump, which is $\frac{\Delta}{\Delta t^h(x)} + o(\Delta)$) from which the martingale property of $B^h(\cdot)$ follows. The quadratic covariation formula follows from

$$\begin{aligned}
&E[(B^h(t+\Delta) - B^h(t))(B^h(t+\Delta) - B^h(t))' | \mathscr{B}^h(t), \xi^h(t) = x] \\
&= E[(\xi^h(t+\Delta) - \xi^h(t))(\xi^h(t+\Delta) - \xi^h(t))' | \mathscr{B}^h(t), \xi^h(t) = x] + o(\Delta) \\
&= \bar\Sigma_h(x)\Delta t^h(x)\left(\frac{\Delta}{\Delta t^h(x)}\right) + o(\Delta).
\end{aligned}$$

Part 2. The tightness of $F^h(\cdot)$ is obvious, together with the fact that any weak limit must be continuous w.p.1. The proof for $\{B^h(\cdot)\}$ closely follows that of [3], Theorem 6.3.1. In particular, if there is a real K such that, for each $T > S$,

$$E|B^h(T) - B^h(S)|^4 \leq K(T-S)^2 + O(h), \tag{8.6}$$

then, using the martingale property of $B^h(\cdot)$, we get both tightness and continuity of all limits, as in [3], Chapter 6. For notational convenience, suppose that $B^h(\cdot)$ is scalar valued.

We will evaluate

$$E \int \int_S^T \int \int \prod_{i=1}^4 dB^h(s_i)$$

by evaluating the four quantities (a) - (d). The l.h.s. of (8.6) is bounded above by a constant times the sum of (a) - (d). Note that $B^h(\cdot)$ has an absolutely continuous component, and a pure jump component.

(a) $E \int_S^T dB^h(t) \int \int_S^t \prod_{i=2}^4 dB^h(s_i)$

(b) $E \int_S^T [dB^h(t)]^2 \int_S^t \int \prod_{i=3}^4 dB^h(s_i)$

(c) $E \int_S^T [dB^h(t)]^3 \int_S^t dB^h(s)$

(d) $E\int_S^T [dB^h(t)]^4$.

(d) is clearly bounded above by some Kh^4 times the average number of jumps in $[S,T]$, which is less than $(T-S)K/h^2$, for some real K. The quantity (a) is zero, by the orthogonality of the increments of $B^h(\cdot)$. The quantity (b) equals, for some real K,

$$E\int_S^T \overline{\Sigma}_h(\xi^h(t))dt \int_S^t \prod_{i=1}^{2} dB^h(s_i)$$

$$\leq K(T-S)\ E\int_S^T \int \prod_{i=1}^{2} dB^h(s_i)$$

$$\leq K^2(T-S)^2.$$

Quantity (c) is bounded above by

$$E\int_S^T (dB^h(t))^2[(dB^h(t))^2 + (\int_S^t dB^h(s))^2],$$

which we can bound by using the results for (b) and (d).

Putting all the estimates together yields (8.6) for some real K.

Thus, $\{\xi^h(\cdot), B^h(\cdot)\}$ is tight. Let h index a convergent subsequence with limit $\xi(\cdot), B(\cdot)$. By the method of Theorem 6.3.2 of [3], it can be shown that $B(\cdot)$ is a

continuous martingale with quadratic covariation $\int_0^t 2a(\xi(s))ds$, and that (8.4) holds for a Wiener process with the asserted properties.

Part 3. The proof of the assertion concerning the representation (8.5) is close to that given in [3], Chapter 6.6, for a similar result. Let $\psi^h(\cdot)$ denote a standard R^r valued Wiener process, which is independent of $\xi^h(\cdot)$. Choose measurable diagonal $D_h(\cdot)$ and orthonormal $P_h(\cdot)$ matrices such that $\bar{\Sigma}_h(x) = \bar{P}_h(x)\bar{D}_h^2(x)\bar{P}_h'(x)$, where $\bar{\sigma}_h(x) = \bar{P}_h(x)\bar{D}_h(x) \to \sigma(x)$ as $h \to 0$. Define $\Sigma_h(t) = \bar{\Sigma}_h(\xi^h(t))$, $P^h(t) = \bar{P}^h(\xi^h(t))$ and $D_h(t) = \bar{D}_h(\xi^h(t))$.

Let $(d_1(t),\ldots,d_r(t))$ denote the diagonal elements of $D_h(t)$. Choose $\alpha \in (0,1)$. Define the diagonal matrices $D_h^+(t), D_h^{++}(t), D_h^T(t)$, resp., as the matrices with i^{th} diagonal elements $d_i^{-1}(t)I_{\{d_i(t)>0\}}$, $d_i^{-1}(t)I_{\{d_i(t)\geq h^\alpha\}}$ and $d_i(t)I_{\{d_i(t)\geq h^\alpha\}}$, resp.

Define $W^h(t)$ by $W^h(0) = 0$ and

$$dW^h(t) = D_h^{++}(t)P_h'(t)dB^h(t) + (I-D_h^T(t)D_h^{++}(t))d\psi^h(t).$$

Then (8.5) follows from[+]

$$dB^h(t) = P_h(t)D_h(t)dW^h(t) + dE^h(t) = \sigma_h(\xi^h(t))dW^h(t) + dE^h(t),$$

[+] $dE^h(t)$ is defined analogously to the ε_n^h above (6.6.6) in [3].

where $E^h(\cdot)$ is a process which tends weakly to the zero process as $h \to 0$. The assertions concerning the tightness of $\{W^h(\cdot)\}$, the non-anticipative property of $W(\cdot)$, and the representation $\int_0^t \sigma(\xi(s))dW(s)$ are all similar to the proofs of the related assertions in [3], Chapter 6.6, and are omitted. Q.E.D.

In [3], Chapters 8 and 9, to prove optimality of the limit of the costs for the discretized problems, it was frequently necessary to "discretize" an arbitrary control for the continuous process $X(\cdot)$, and then to apply this to the discrete model. The $W^h(\cdot)$ obtained in this Section can be used instead of the $W^h(\cdot)$ of those theorems (e.g., [3] Theorem 8.2.4, etc.)

8.2. <u>The General Case</u> ($g, Q \not\equiv 0$)

Of the several ways in which a Markov $\xi^h(\cdot)$ can be defined, we will develop one that is particularly easy to relate to formulas such as (2.4), (2.5) and (3.5). Pure jump right continuous processes $A^h(\cdot), J^h(\cdot)$ and $\xi^h(\cdot)$ will be defined such that

$$\xi^h(t) - \xi^h(s) = \int_s^t dA^h(u) + \int_s^t dJ^h(u) \tag{8.7}$$

<u>Define</u> $\overline{\Delta t}^h(x) = h^2/Q_h(x)$. <u>Note that</u> $\overline{\Delta t}^h(x)$ <u>is</u>

defined as $\Delta t^h(x)$ was in Section 8.1, but not as $\Delta t^h(x)$ was in Sections 2 to 7. The processes are defined by the following relations. Let $S(t)$ denote the set $\{\xi^h(s), A^h(s), J^h(s), s \leq t\}$. All $o(\cdot)$ are uniform in ω, t, but not necessarily in h.

(8.8a) $P\{2 \text{ or more jumps of } \{A^h(\cdot), J^h(\cdot)\} \text{ in } (t,t+\Delta] \,|\, S(t)\} = o(\Delta)$,

(8.8b) $P\{A^h(\cdot) \text{ jumps in } (t,t+\Delta] \,|\, S(t), \xi^h(t) = x\} = \Delta/\overline{\Delta t}^h(x) + o(\Delta)$,

(8.8c) $P\{J^h(\cdot) \text{ jumps in } (t,t+\Delta] \,|\, S(t)\} = c\Delta + o(\Delta)$.

Let τ_i and σ_i denote the i^{th} jump times of $A^h(\cdot)$ and $J^h(\cdot)$, resp. Then let

(8.8d) $P\{A^h(\tau_i) - A^h(\tau_i^-) = y \,|\, S(\tau_i^-), \tau_i, \xi^h(\tau_i^-) = x\} = p^h(x,y)$,

(8.8e) $P\{J^h(\sigma_i) - J^h(\sigma_i^-) = (j_1'h, \ldots, j_r'h) \,|\, S(\sigma_i^-), \sigma_i, \xi^h(\sigma_i^-) = x\} =$

$$= \sum_{(j_1,\ldots,j_r \text{ corresp. to } j_1',\ldots,j_r')} \Gamma^h(x, j_1, \ldots, j_r),$$

(see Section 2). Note that (8.8e) allows jumps of $J^h(\cdot)$ of zero magnitude, as in Sections 2 to 4.

Suppose that $\xi^h(\cdot)$ jumps at t, with $\xi^h(t) = \xi^h(t^+) = x$. Then $\xi^h(\cdot)$ is constant until the next jump. The interval until the next jump is the smallest of two random variables, one exponentially distributed with mean $1/c$, the second exponentially distributed with mean $\overline{\Delta t}^h(x)$. Thus, given $S(t)$ with $\xi^h(t) = x$, the probability that the next jump after t is that of $J^h(\cdot)$, rather than that of $A^h(\cdot)$, is

$$\frac{\overline{\Delta t}^h(x)c}{1+\overline{\Delta t}^h(x)c}, \tag{8.9}$$

which is precisely $\hat{p}^h(x)$ (see (2.3a)). Thus, the probability that $J^h(\cdot)$ jump next is just the probability that the change in $\xi^h_{n+1} - \xi^h_n$ (given $\{\xi^h_n = x\}$ and using (2.3a)) is due to the "jump component" of the process $\{\xi^h_i\}$.

The average (conditioned on $S(t), \xi^h(t) = x$) time until the next jump after t of $\xi^h(\cdot)$ is

$$\frac{\overline{\Delta t}^h(x)}{c\overline{\Delta t}^h(x)+1} \tag{8.10}$$

which is precisely the $\Delta t^h(x)$ of Section 2. Let $\{u_i\}$ denote the jump times of $\xi^h(\cdot)$. Then the distribution of $\xi^h(u_{i+1}) - \xi^h(u_i)$, given $\xi^h(u_i) = x$, is the same as that of $\xi^h_{n+1} - \xi^h_n$ given $\xi^h_n = x$. This, together with the expression for the conditional average waiting time (8.10), implies that (3.5) remains the solution to (2.4), where $\xi^h(\cdot)$ is the process just constructed, and ρ_h its' escape time from G_h.

Thus, the new interpolation makes sense, for our purposes. It is clearly a Markov jump process, and can be used to study the limit of $V^h(x)$, and of other functionals of the chain, provided that $\xi^h(\cdot) \to X(\cdot)$ in distribution, as $h \to 0$, and (A1.3) holds.

Using the definition of $B^h(\cdot)$ and $F^h(\cdot)$ of Section 8.1, and letting $\xi^h(0) = x$,

$$\xi^h(t) = x + \int_0^t f(\xi^h(s))ds + B^h(t) + J^h(t) \tag{8.11}$$

Let $\hat{Q}^h(\cdot,\cdot)$ denote a Poisson measure (with corresponding Poisson process $\hat{Q}^h(\cdot)$) satisfying (A1.2). Then we can assume that $J^h(\cdot)$ has the representation

$$J^h(t) = \int_0^t \int g(\xi^h(s^-),\alpha)\hat{Q}^h(d\alpha \times ds) + \hat{E}^h(t), \tag{8.12}$$

where $\hat{E}^h(\cdot) \to$ zero process weakly, as $h \to 0$, and $\xi^h(\cdot)$ is non-anticipative with respect to $\hat{Q}^h(\cdot)$. The $\hat{E}^h(\cdot)$ represents only the indefinite sums of the difference between the jumps in the integral in (8.12) and the points on the grid G_h to which our convention assigns these points; e.g., the assigned point is $(j_1'h,\ldots,j_r'h)$, if the jump is in the box $b^h(j_1,\ldots,j_r)$. We can assume the form (8.12), in the sense that there are processes $\xi^h(\cdot),\hat{E}^h(\cdot),A^h(\cdot)$ satisfying the properties below (8.12) and such that (8.3) continues to hold and where $\xi^h(\cdot)$ is defined by

$$\xi^h(t) = x + A^h(t) + \int_0^t \int g(\xi^h(s^-),\alpha)\hat{Q}^h(d\alpha\times ds) + \hat{E}^h(t),$$

(8.13)

$$A^h(t) = F^h(t) + B^h(t),$$

and the $(\xi^h(\cdot),F^h(\cdot),B^h(\cdot),J^h(\cdot) + \hat{E}^h(\cdot))$ in (8.13) induce the same measure on $D^{4r}[0,\infty)$ as does the set in (8.11) (with $\hat{E}^h(\cdot)$ deleted). The only purpose of $\hat{Q}^h(\cdot,\cdot)$ is to generate the jumps of $J^h(\cdot)$, according to the appropriate distribution. Define $W^h(\cdot)$ as in Section 8.1.

Theorem 8.2. Assume (A1.1) - (A1.2). Define $\xi^h(\cdot)$, $F^h(\cdot),B^h(\cdot),J^h(\cdot)$ (and $\hat{E}^h(\cdot),\hat{Q}^h(\cdot)$, where appropriate) by the construction leading to either (8.11) or (8.13). The assertions of Theorem 8.1 concerning $B^h(\cdot)$ and $W^h(\cdot)$ continue to hold. The processes above are all tight, and $\hat{E}^h(\cdot) \to$ zero process, $\hat{Q}^h(\cdot) \to$ Poisson process $\hat{Q}(\cdot)$, weakly as $h \to 0$. Let h index a convergent subsequence of the other processes, with limit $\xi(\cdot),F(\cdot),B(\cdot),J(\cdot)$. There exists a Wiener process $W(\cdot)$ (which can be the limit of $W^h(\cdot)$) such that

$$\xi(t) = x + \int_0^t f(\xi(s))ds + B(t) + J(t),$$

where

$$B(t) = \int_0^t \sigma(\xi(s))dW(s),$$

and $\xi(\cdot),B(\cdot),J(\cdot)$ (and $\hat{Q}(\cdot)$, where appropriate) are non-anticipative with respect to $W(\cdot)$. Indeed, $\hat{Q}(\cdot)$ and $W(\cdot)$ are independent.

The process $J(\cdot)$ has the same properties as the $J(\cdot)$ in Section 4. In the case of the construction (8.13) and if $g(\cdot,\cdot)$ is continuous,

$$J(t) = \int_0^t \int g(\xi(s^-),\alpha)\hat{Q}(d\alpha \times ds), \text{ all } t, \text{ w.p.1.}$$

Under (A1.3), the limit $\xi(\cdot)$, under either construction, induces the same measure on $D^r[0,\infty)$ as does the solution of (1.1).

The proof is a combination of those of Theorems 4.1, 4.2, 5.2 and 8.1 and is omitted.

Under (A1.3) and (A1.4), the processes of this section and that of Section 3 are asymptotically weakly equivalent to (1.1). The process of this section is clearly Markov. The results of Section 7 hold without (A1.4). The process in (3.5) can be replaced by that of (8.11) or (8.13). Equation (3.5) holds for the interpolation of Section 3 even if (A1.4) does not hold, but in that case, the weak limits of those interpolations are not necessarily equivalent to (1.1).

9. Optimal Stoppping Problems and Auxiliary Results

First, we prove a result which is useful to show that a control or stopping time is suitably non-anticipative. Let $\beta(\cdot), Q(\cdot), w(\cdot)$ and Y be processes with paths in some $D^P[0,\infty)$, and a random variable, such that $Q(\cdot)$ is a Poisson process with the properties of (A1.1), $w(\cdot)$ is a Wiener process and, for each integer b and bounded continuous $f_i(\cdot)$ and each $t \geq s$ (a given number) and real numbers $t_1,\ldots,t_b \leq t$, let

$$(9.1) \quad Ef_1(Y)f_2(Q(t_i),w(t_i),\rho(t_i),\ i \leq b)(w(t+u) - w(t)) = 0, \quad u \geq 0,$$

$$(9.2) \quad Ef_1(Y)f_2(Q(t_i),w(t_i),\rho(t_i),\ i \leq b)\ A(t,u) = 0,\ u \geq 0,$$

$$A(t,u) = (w(t+u) - w(t))(w(t+u) - w(t))' - uI$$

$$(9.3) \quad \{Q(s+u) - Q(s),\ u \geq 0\} \text{ is independent of } \{Q(u),w(u),\rho(u),\ u \leq s,Y\}.$$

Theorem 9.1. Under the above assumptions with $s = 0$, $Q(\cdot), w(\cdot)$ have independent increments. In general $\{Q(s+u) - Q(s),\ w(s+u) - w(u),\ u \geq 0\}$ is independent of $\{Q(u),w(u),\rho(u),\ u \leq s,Y\}$.

Proof. Since the first assertion follows from the second, only the second will be proved. Let $\sigma_0 = s$ and

$\{\sigma_j, j \geq 1\}$ denote the jump times of $Q(\cdot)$ on $(s,T]$, where T is an arbitrary number $> s$. Let $\lambda(\cdot)$ and $\eta(\cdot)$ denote vector valued bounded continuous functions on $[0,T]$, and let v denote a vector. Define

$$A(t) = \exp i[v'Y + \int_0^t \lambda'(u)dW(u) + \int_0^t \eta'(u)dQ(u)].$$

By the assumption (9.1) - (9.2), $w(\cdot)$ is a martingale and a Wiener process on $[s,T]$ with respect to $\{\mathscr{G}(t), t \geq s\}$, where $\mathscr{G}(t) = \mathscr{B}(Y,w(u),Q(u),\rho(u), u \leq t)$. Also, the σ_j are stopping times with respect to $\{\mathscr{G}(t), t \geq s\}$.

We have

$$A(t) = A(s) + \sum_{\substack{j\geq 1\\ \sigma_j\leq t}} [A(\sigma_j^-) - A(\sigma_{j-1})] + \sum_{\substack{j\geq 1\\ \sigma_j\leq t}} [A(\sigma_j) - A(\sigma_j^-)].$$

By Itô's Lemma, the first sum is $(A(u^-)$ can be replaced by $A(u)$ here)

$$\sum_{\substack{j\geq 1\\ \sigma_j\leq t}} \int_{\sigma_{j-1}}^{\sigma_j} A(u^-)[i\lambda'(u)dw(u) - \frac{|\lambda(u)|^2}{2} du].$$

The second sum is

$$\sum_{\substack{j\geq 1\\ \sigma_j\leq t}} A(\sigma_j^-)[(\exp i\eta'(\sigma_j)dQ(\sigma_j)) - 1]$$

$$= \int_s^t (\exp i[v'Y + \int_0^t \lambda'(u)dw(u)]d_t[\exp i\int_0^t \eta'(u)dQ(u)].$$

Define $E(t) = EA(t)$. Then, (9.1) - (9.3) imply that

$$E(t) = E(s) + \int_s^T E(u)[- \frac{|\lambda(u)|^2}{2} + c(\mathscr{U}(u) - 1)]du, \tag{9.4}$$

where

$$\mathscr{U}(t) = E \exp i\, \eta'(t)X,$$

where X has the distribution of $Q(\sigma_j) - Q(\sigma_j^-)$, the jump in $Q(\cdot)$. The unique solution to (9.4) is

$$E(t) = E(s)\exp[\int_s^t \frac{|\lambda(u)|^2}{2}\, du + \int_s^t c(\mathscr{U}(u) - 1)du]. \tag{9.5}$$

The result follows from the product form of the characteristic functional in (9.5), since the exponential term is the characteristic functional of $\{w(s+u) - w(s), Q(s+u) - Q(s), 0 \leq u \leq t - s\}$. Q.E.D.

<u>Applications of Theorem 9.1</u>. Let $\hat{Q}(\cdot), W(\cdot)$ denote the limits in the past sections. Set $s = 0$, and drop Y, set $\hat{Q}(\cdot) = Q(\cdot)$, $W(\cdot) = w(\cdot)$, $\xi(\cdot) = \rho(\cdot)$. Equations (9.1) - (9.3), and the assumptions

above them, follow by the type of weak convergence arguments used in the proof of [3], Theorems 6.3.2 and 6.6.1. In Theorem 6.3.2, the $\xi(s_i)$ (with or without the superscript h) are to be replaced by $\xi(s_i), Q(s_i), W(s_i)$ (with or without the superscript h).

Theorem 9.1 is the analog for our case of [3], Theorem 8.2.1 and Corollary 8.2.1, which are used to show that a stopping time ρ which is a limit of $\{\rho_h\}$, where ρ_h is an optimal stopping time for an optimal stopping problem on the chain, is non-anticipative with respect to $W(\cdot)$. Here, we would let $Y = I_{\{\rho \le s\}}$.

Theorem 9.1 is the analog - for our jump case-of [3] Theorem 8.2.1 and Corollary 8.2.1. These theorems were used to show that certain limits of optimal controls or stopping times for optimal control or stopping problems on controlled forms of $\{\xi_n^h\}$, actually converged to non-anticipative times and controls for the limiting controlled process. The same thing is done here. For example, in the context of [3], Corollary 8.2.1, we would set $Y = I_{\{\rho \le s\}}$, where ρ is the limit of the "approximate" stopping times, and show (9.1) - (9.3) by weak convergence arguments as in [3].

The technique in [3], for proving optimality of a limiting control or stopping time, involved comparing the cost under the optimal policy for each controlled chain $\{\xi_n^h\}$ to the cost under a "discretized" form of a policy for the controlled diffusion. The same thing is done here. We only

illustrate the analog of the technique of [3], Theorem 8.2.4, to which the reader is referred. The remarks below will suggest the necessary alterations to the other policy discretization techniques of [3], Chapters 8 and 9.

Let $\tau \leq T$ denote a stopping time for (1.1) which is a functional of $w(\cdot), Q(\cdot)$. Let $0 < \delta << \Delta < T$ with $\Delta/\delta = q$, an integer. Let $\mathcal{B}^{\delta}_{n\Delta}(w,Q)$ denote the smallest σ-algebra which measures $\{w(i\delta), Q(i\delta),\ i\delta \leq n\Delta\}$. Define

$$\rho_i(\delta,\Delta) = E_x[\tau \mid \mathcal{B}^{\delta}_{i\Delta}(w,Q)],\quad i \geq 1,$$

and

$$\tau(\delta,\Delta) = \sum_{i=1}^{T/\Delta} (i\Delta) I_{\{\rho_i(\delta,\Delta) \leq i\Delta,\ \rho_j(\delta,\Delta) > j\delta,\ j<i\}}.$$

If $\tau(\delta,\Delta)$, or any other stopping time, is not defined at some ω, set it equal to T there. Let the decision sets for $\tau(\delta,\Delta)$ be $A_i(\delta,\Delta) \in R^{iq(r+r')}$ (i.e., $\tau(\delta,\Delta) = i$ if $\{w(j\delta), Q(j\delta),\ j\delta \leq i\Delta\} \in A_i(\delta,\Delta)$).

For each $\varepsilon > 0$, there are sets $A^{\varepsilon}_i(\delta,\Delta)$ and $A_{\varepsilon,i}(\delta,\Delta)$, which are open and closed, resp., and which satisfy:

$$A^{\varepsilon}_i(\delta,\Delta) \supset A_i(\delta,\Delta) \supset A_{\varepsilon,i}(\delta,\Delta),$$

$$\sum_i P\{(w(j\delta), Q(j\delta),\ j\delta \leq i\Delta) \in A^{\varepsilon}_i(\delta,\Delta) - A_{\varepsilon,i}(\delta,\Delta)\} \leq \varepsilon.$$

For each i, there is an open set, $\tilde{A}_i^\varepsilon(\delta,\Delta)$, containing $A_{\varepsilon,i}(\delta,\Delta)$, and whose closure is in $A_i^\varepsilon(\delta,\Delta)$, and such that

$$P\{(w(j\delta),Q(j\delta),j\delta \le i\Delta) \;\varepsilon\; \partial\tilde{A}_i^\varepsilon(\delta,\Delta)\} = 0.$$

These sets $\{\tilde{A}_i^\varepsilon(\delta,\Delta)\}$ are used to get the approximate stopping times, just as the B_i^n were in the reference. Define

$$C_i^\varepsilon = \{v \;\varepsilon\; R^{iq(r+r')} : v \;\varepsilon\; \tilde{A}_i^\varepsilon(\delta,\Delta),\; v \not\varepsilon\; \tilde{A}_j^\varepsilon(\delta,\Delta) \times R^{(i-j)q(r+r')},\; j < i\}.$$

Define

$$\begin{aligned}\tau^\varepsilon(\delta,\Delta) &= i\Delta \quad \text{if} \quad \{w(j\delta),Q(j\delta),\; j\delta \le i\Delta\} \;\varepsilon\; C_i^\varepsilon\,, \\ &= T \quad \text{if not otherwise defined.}\end{aligned}$$

The $\tau^\varepsilon(\delta,\Delta)$ with $(W^h(j\delta),\hat{Q}^h(j\delta))$ substituted for $(w(j\delta),Q(j\delta))$ is the appropriate discretized comparison stopping time for $\{\xi_n^h\}$. The rest of the development is as in the reference, except that $\varepsilon \to 0$ here replaces $n \to \infty$ there.

REFERENCES

[1] I.I. Gikhman, A.V. Skorokhod, Stochastic Differential Equations, Springer-Verlag, New York, 1972.

[2] H.J. Kushner, "A Survey of Some Applications of Probability and Stochastic Control Theory to Finite Difference Methods for Elliptic and Parabolic Equations", SIAM Review 18, pp. 545-577, 1976.

[3] H.J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, New York, 1976 or early 1977.

[4] H.J. Kushner, C.F. Yu, "Probability Methods for the Convergence of Finite Difference Methods to Partial-Differential-Integral Equations", J. Math. Anal. and Applic., 45, pp. 54-72, 1974.

[5] I.I. Gikhman, A.V. Skorokhod, Introduction to the Theory of Random Processes, Saunders, Philadelphia, 1969.

[6] P. Billingsley, Convergence of Probability Measures, John Wiley, New York, 1968.

SYMBOL LIST

$\int$	large integral
τ	tau
π	pi (lower case)
ϕ	phi (lower case)
Γ	gamma (upper case)
{ }	curly braces
ε	epsilon (lower case)
$\times$	times
$\sum$	sigma (upper case)
∂	differential
∞	infinity
[ ]	brackets (square)
δ	delta (lower case)
θ	theta (lower case)
ρ	rho (lower case)
ψ	psi (lower case)
η	eta (lower case)

σ	sigma (lower case)
$\equiv$	identical to
$<$	less than
$>$	greater than
$\cap$	intersection
Π	pi (upper case)
$\pm$	plus or minus
Δ	delta (upper case)
ξ	xi (upper case)
^	carat
~	similar cycle
β	beta (lower case)
Φ	phi (upper case)
ω	omega (lower case)
λ	lambda (lower case)
$\vert$	single parallel bar

<u>Script</u>

$\mathscr{D}$	D
$\mathscr{B}$	B
$\mathscr{L}$	L
$\mathscr{S}$	S
$\mathscr{U}$	U

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
AFOSR - TR - 77 - 0018

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
APPROXIMATIONS FOR FUNCTIONALS AND OPTIMAL CONTROL PROBLEMS ON JUMP DIFFUSION PROCESSES

5. TYPE OF REPORT & PERIOD COVERED
Interim

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Harold J. Kushner and Giovanni DiMasi

8. CONTRACT OR GRANT NUMBER(s)
~~XXXXXXXXXX~~
AFOSR 76-3063

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University, Providence, R.I. 02912

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
2304/A1
61102F

11. CONTROLLING OFFICE NAME AND ADDRESS
Air Force Office of Scientific Research
Bolling Air Force Base (NM)
Washington, D.C.

12. REPORT DATE
November 1976

13. NUMBER OF PAGES
54

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper treats approximations to stochastic differential equations with both a diffusion a-nd a jump component, and to associated functionals and partial-differential-integral equations of the (degenerate or not) elliptic or parabolic type. Approximations for the optimal control problem on such a morel, or for the associated non-linear partial-differential-integral equation is discussed. The techniques are purely probabilistic and are extensions of those in [3], which dealt with the diffusion case.